AD_____

AWARD NUMBER: W81XWH-07-1-0325

TITLE: Multiplex Quantitative Histologic Analysis of Human Breast Cancer Cell Signaling and Cell Fate

PRINCIPAL INVESTIGATOR: William M. F. Lee; Badrinath Roysam

CONTRACTING ORGANIZATION: University of Pennsylvania
Philadelphia, PA 19104

REPORT DATE: May 2009

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
1 May 2009	Annual	1 May 2008 – 30 Apr 2009

4. TITLE AND SUBTITLE

Multiplex Quantitative Histologic Analysis of Human Breast Cancer Cell Signaling and Cell Fate

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-07-1-0325

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

William M. F. Lee; Badrinath Roysam

E-Mail: leemingf@mail.med.upenn.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Pennsylvania
Philadelphia, PA 19104

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Many molecular events and cellular processes are preserved in fixed human tumor specimens, and access to this information awaits a method for them to be quantified and analyzed. Hormone receptors, HER2, cell signaling and proliferation events are prognostically and therapeutically important in human breast cancer and can be revealed by immunohistological staining. A novel platform for study of immunostained breast cancer specimens is being developed that will quantify analyte antigens on a cellular basis, i.e. cytometrically. The platform uses multispectral microscopy to examine breast cancer specimens that have been immunostained for multiple cell type and analyte antigens using different chromogens and fluorophores. Multispectral microscopy and spectral separation permits staining for individual antigens to be distinguished and separated from staining for other antigens in multiplex-stained slides. Stains for structural and cell type antigens (e.g. nuclei, epithelial cytokeratins, E-cadherin) are used by FARSIGHT software to segment individual nuclei and cells in images and to identify those that are breast cancer cells. Stains for biomarkers and cell signaling antigens (e.g. ER, PR, HER2. Ki67, p-ERK, p-AKT) are then associated with the segmented cells to quantify expression of these analytes in breast cancer cells on a per-cell basis. Progress in the past year has developed the cell segmentation function of FARSIGHT software so that nuclear (ER, PR, Ki67) and cytosolic/cell membrane analytes (HER2) can be cytometrically quantified in human breast cancer specimens.

15. SUBJECT TERMS
Breast cancer, cell signaling, cell proliferation, histology, image analysis

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	11	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Principal Investigator: **Lee, William M. F.**

Proposal Title: **Multiplex Quantitative Histologic Analysis of Human Breast Cancer Cell Signaling and Cell Fate**

Introduction

The objective of this proposal is to build an advanced platform for immunohistological study of breast cancer specimens that retrieves multiplex quantitative molecular information on a cellular basis. Three components make up this platform: (1) Multiplex tissue immunostaining protocols for revealing structural, cell-type and analyte antigens in the same histologic section (the first two to help segment and type cells, and the last to reveal biological processes/events and prognostic/predictive biomarkers of interest); (2) imaging by multispectral microscopy to capture the information revealed by individual stains in multiplex combinations; and (3) software (FARSIGHT) for automated multispectral image analysis that (i) segments individual nuclei and cells in images, (ii) classifies the segmented nuclei/cells into cell types of interest based on spatial/ textural features and their association with structural/cell-type antigen staining, and (iii) quantifies analyte expression on a cellular basis by associating analyte staining with the segmented nuclei/cells. Operation of the platform will be developed using human breast cancer specimens with the goal of quantifying histological prognostic/predictive biomarkers (ER, PR, HER2) and biological events/processes (Ki67, p-ERK) cytometrically.

We originally proposed to apply the analytical system to analyze the pharmacodynamics of lapatinib in tumor samples from patients in a clinical trial of neoadjuvant lapatinib (PI, Dr. Angela DeMichele). We planned to compare breast cancer cell expression of Ki67, p-ERK, p-AKT and HER2 in core tumor biopsies obtained before lapatinib was started and after two weeks of lapatinib therapy. This trial was stopped after two years because of a lack of patient accrual - only two patients were ever recruited. Accordingly, this goal has been removed from our project.

Task 1. Develop robust protocols for multiplex immunostaining of human breast cancer specimens

We have developed robust immunostaining protocols for detecting analyte antigens in paraffin-embedded human breast cancer specimens that report on cell signaling events (p-ERK, p-AKT, p-STAT3), cell fate decisions (Ki-67) and biomarkers of prognostic and predictive value (ER, PR, HER2). Immunohistochemical (IHC) staining for these analytes using chromogenic substrates (DAB, diaminobenzidine) consistently yielded the highest percentage of positive tumor cells (with low background staining). Staining using fluorescent reporters yielded comparable results for some of the analytes (p-ERK, Ki-67, HER2), making immunofluorescent (IF) staining an option when studying these analytes. For other analytes (e.g. p-AKT, p-STAT3), however, IF staining yielded far fewer positive cells, indicating that IF was signficantly less sensitive than IHC for studying these analytes. Structural and cell-type antigens (epithelial cytokeratin, CK, E-cadherin, HER2) used for typing segmented nuclei and cells as breast carcinoma cells are generally abundant and equally well revealed by IHC or by IF immunostaining. Putting these individual antigen stains together in multiplex immunostaining protocols, we propose to use IHC staining for the single analyte stained on a slide (employing DAB chromogen) followed by IF staining for the structural/cell-type antigens on the slide. The exception is when HER2 is our analyte, in which case we use IF staining for both HER2 analyte and CK cell-type/structural antigen. To prevent nonspecific staining by the secondary anti-mouse Ig antibody used in IF staining due to cross-binding to mouse antibodies applied previously during analyte IHC staining, we incubate slides following DAB staining in a 5% SDS solution at 50^{o}C for 5 minutes. This “strips” off all antibodies present on the slide so that these cannot cause nonspecific staining in subsequent IF immunostaining steps.

To study two analytes in the same tumor section (e.g. p-ERK+Ki67 or ER+PR), we developed dual analyte IHC staining protocols using DAB as the reporter for one analyte and SG Blue as the reporter for the other analyte. Staining for cell-type/structural antigens follows completion of IHC analyte

immunostaining. Nuance multispectral microscopy is used to detect and separate staining by the two chromogens, even when they are co-localized.

We have determined that phospho-epitopes (p-ERK, p-AKT, p-STAT3) are lost over time in cut tumor sections. Significant loss of these epitopes is seen in slides a month after sectioning. We have made it a policy to immunostain slides for phosphopeptide analytes within two weeks of sectioning.

Task 2. Optimize multispectral imaging and data capture for subsequent computational analysis

We have tested various image capture conditions and parameters to determine the optimal for acquiring images for accurate analyte quantification and for adequate sampling of breast cancer specimens. We perform analysis at 400X magnification and usually acquire 150-300 breast cancer cells for analysis from each image. We acquire 5-10 images from each specimen, so 750-3000 tumor cells are analyzed.

Recently, we acquired the Vectra Multispectral Imaging System (CRi) which includes a multispectral microscope with a robotic slide loader and a computer-controlled stage. It includes Inform software which allows automated image acquisition at 200X magnification selected from regions of interest (ROI) determined at 40X magnification. The software has the ability to identify regions of breast carcinoma cells in hematoxylin-stained images, and this is used to select the ROI at low magnification for sampling at higher magnification. We will incorporate this imaging system into our analysis platform as it offers the potential to type cells as breast carcinoma cells without immunostaining.

Task 3. Develop computational algorithms for multispectral immunohistological image analysis

A software tool named FARSIGHT (Figure 1) has been developed to quantify intrinsic and associative parameters associated with whole cell segmentation from multi-spectral image data recorded at the PI's laboratory at U. Penn. The system has been implemented using the IDL software development platform (ITT Industries), and is currently undergoing evaluation at the PI's laboratory. Starting with a batch of multiplexed immunostained specimens that are imaged multi-spectrally and unmixed as described above, the software is able to quantify key breast cancer molecular biomarkers (ER, PR, Ki67, and HER2/neu) on a cell-by-cell basis. Each cell in the field is segmented and identified by type, and the intra-cellular distribution of molecular biomarkers is quantified for each cell. A comprehensive list of intrinsic morphological parameters of all cells and associative measurements of the biomarkers are computed from the segmentation. These measurements can be used in addition to pixel-level measurements, and are used for identifying cell types, and then quantifying the biomarkers by cell type. We evaluated the methodology for quantifying estrogen receptor, progesterone receptor, human epidermal growth factor receptor 2 (HER2), and Ki67 in breast cancers. Our methods can be extended to study other cellular antigen(s) in cancer histopathology samples, and in other types of histologic examinations. Our methods were validated quantitatively on controlled *in vitro* specimens prefabricated with known ratios of cell types, and visually on breast human specimens.

Figures 2 and 3 visually illustrate the steps in the automated image analysis. Figure 2 shows a sample breast cancer specimen multiplex stained for HER2 by immunofluorescence using Texas red and for cytokeratin by immunofluorescence using Alexa-488 and counterstained with hematoxylin, and imaged multi-spectrally in absorption and fluorescence modes, and unmixed to yield non-overlapping channels. Panel A is a Brightfield image showing hematoxylin staining. Panel B shows the channel containing only the cell nuclei, corresponding to the hematoxylin spectral signature. Panel C shows he channel corresponding to fluorescently-stained cytokeratin indicating cells if epithelial origin. Panel D shows the channel corresponding to fluorescently-stained HER2, a cell-membrane bound molecular biomarker of interest. Panel E is a composite 3-color image of the data with nuclei (red), cytokeratin (green), and HER2 (blue). Panel F shows the spectral signatures used or the unmixing computations, displayed using the color convention of panel E.

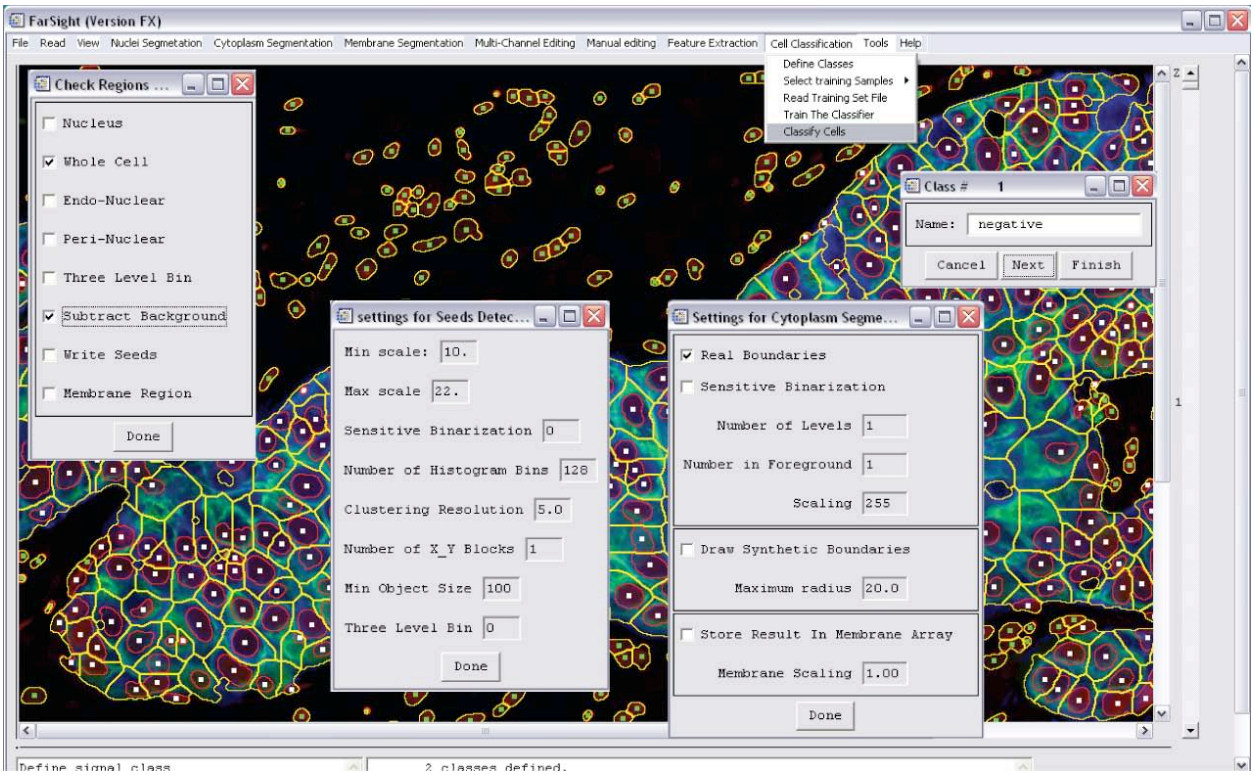

Figure 1: A screen view of the proposed software system built using the IDL programming platform (ITT Corp.). It implements all of the methods proposed, and includes facilities to inspect and edit segmentation results, and export all of the results in tabular form for further analysis at the PI's laboratory. It is currently

Figure 3 illustrates the steps in the automated image analysis for the same sample breast cancer specimen shown in Figure 2, which is labelled for HER2. The first step in the image analysis is automatic segmentation of all cell nuclei. In this regard, we have developed a novel algorithm based on multi-scale Laplacian filtering based algorithms to locate the "marker seed points" shown as green dots in Figure 2A that indicate centers of putative nuclei. The use of these markers greatly simplifies the task of inspecting and editing nuclear segmentation results since the user only needs to add or remove markers. We have found this to be efficient and effective. Once the user approves the marker results, the system automatically delineates the nuclei using a novel graph-cuts segmentation algorithm that significantly improves upon our prior work. In Panels A-E, the nuclear segmentation results are shown as red outlines, and the seeds are shown as green dots. Next, a seeded watershed algorithm is used to delineate the cytosolic boundaries of all cells of epithelial origin. Panel B shows the estimated cytoplasmic domains of the cells of epithelial origin as blue outlines. Panel C shows results of automated segmentation of the cytokeratin channel. This enables us to quantify the amount and intra-cellular distribution of ER, PR, Ki67, and HER2/neu. Panel C also illustrates cell classification results, based on the cytokeratin signal. This image shows the combined nuclear and cytoplsmic segmentation results, with the latter shown in yellow outlines. The yellow dots represent CK+ cells meaning cells of epithelial origin, and pink dots represent CK- cells. Panel D shows the composite cell classification results, with yellow dots indicating cells that are cytokeratin and HER2 positive, and purple dots indicating other cells. Panel E is a close-up view illustrating the regions of interest used to quantify HER2, using red outlines to indicate the nucleus. The region between the blue outline (plasma membrane) and the orange outline (fixed-width spacer) is used to quantify HER2. Panel F is a histogram summary of the cells in this field that were cytokeratin positive: 98.5% of these cells were also HER2 positive (this compares favorably with a manual reading that indicated 99%).

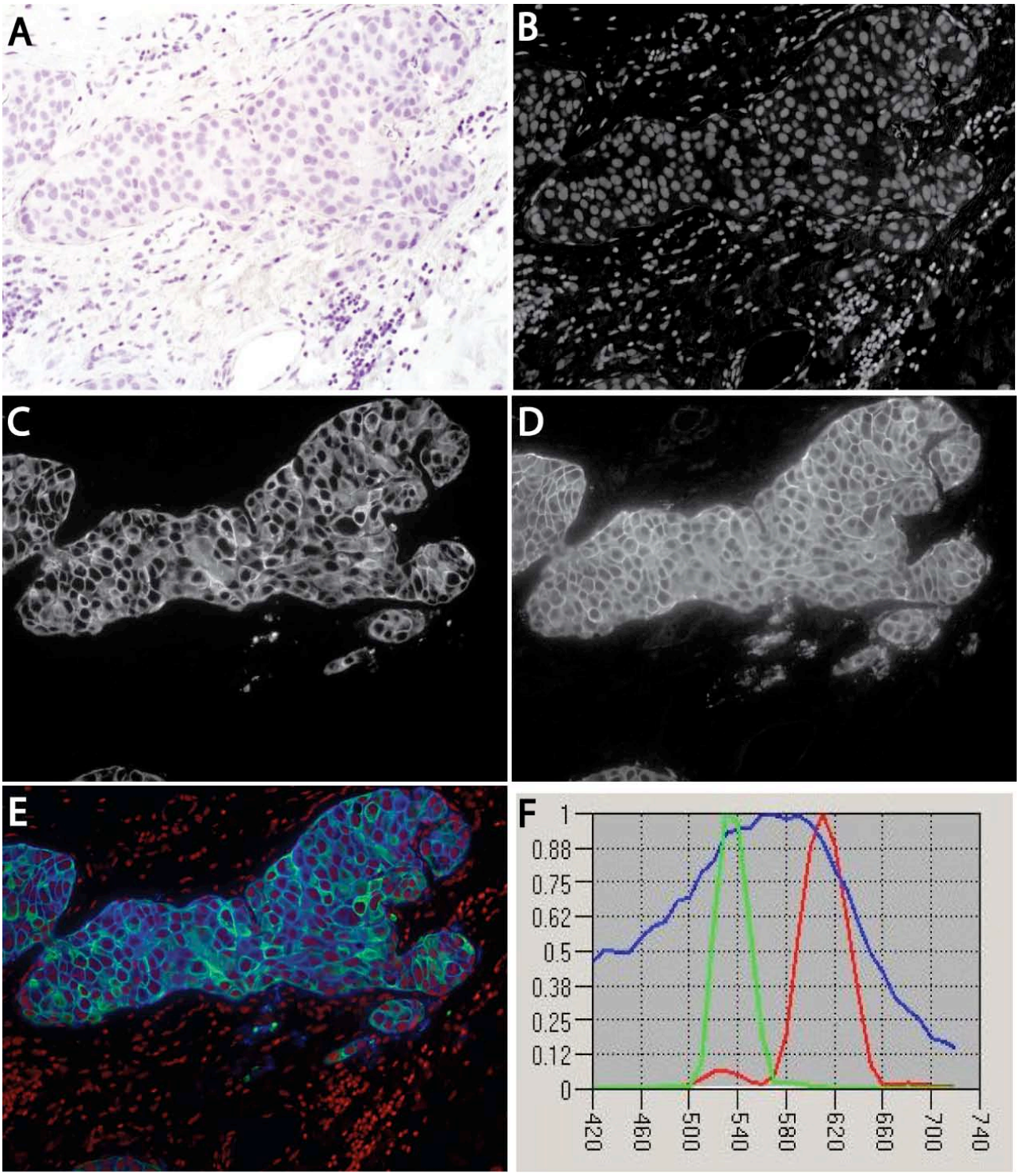

Figure 2: Sample breast cancer specimen multiplex stained for HER2 by immunofluorescence using Texas red and for cytokeratin by immunofluorescence using Alexa-488 and counterstained with hematoxylin, and imaged multi-spectrally in absorption and fluorescence modes, and unmixed to yield non-overlapping channels. (A) Brightfield image showing hematoxylin staining. (B) Channel containing only the cell nuclei, corresponding to the hematoxylin spectral signature. (C) The channel corresponding to fluorescently-stained cytokeratin. (D) The channel corresponding to fluorescently-stained HER2. (E) Composite 3-color image with nuclei (red), cytokeratin (green), and HER2 (blue). (F) Spectral signatures used or the unmixing computations, displayed using the color convention of panel E.

The above methodology was also extended to ER, PR, and Ki67 analysis. The main differences are in the subcellular localization of the markers, and the need to correct the channels recording these biomarkers for a high background level using a 3-level Otsu segmentation algorithm. In each case, the automatically generated results were compared with manual scoring results. An early limited scale study produced the data in table 1 below. The percentages of positive cells were very close to the ground truths, which were estimated by manually marking positive (ER+, PR+ or Ki67+) cells in the images. More specifically, our automatic percentages of ER+, PR+ and Ki67+ were 39%, 40% and 27% of the tumor cells (CK+) respectively. On the other hand, manual percentages were 38%, 39% and 26% respectively. Comparisons between manual and automatic classification results are given in Table 1.

Table 1: Summary of classification results corresponding to the examples shown.

Image	**Percentage of positive cells**	
	Manual Average (%)	**Automated Average (%)**
In_vitro 0%	0	2.4
In_vitro 10%	10	11.1
In_vitro 33%	33	32.8
In_vitro 50%	50	53.7
In_vitro 66%	66	66.9
In_vitro 90%	90	86.4
In_vitro 100%	100	98.0
In_vivo ER	38	39
In_vivo PR	39	40
In_vivo Ki67	26	27
In_vivo HER2	99	98.5

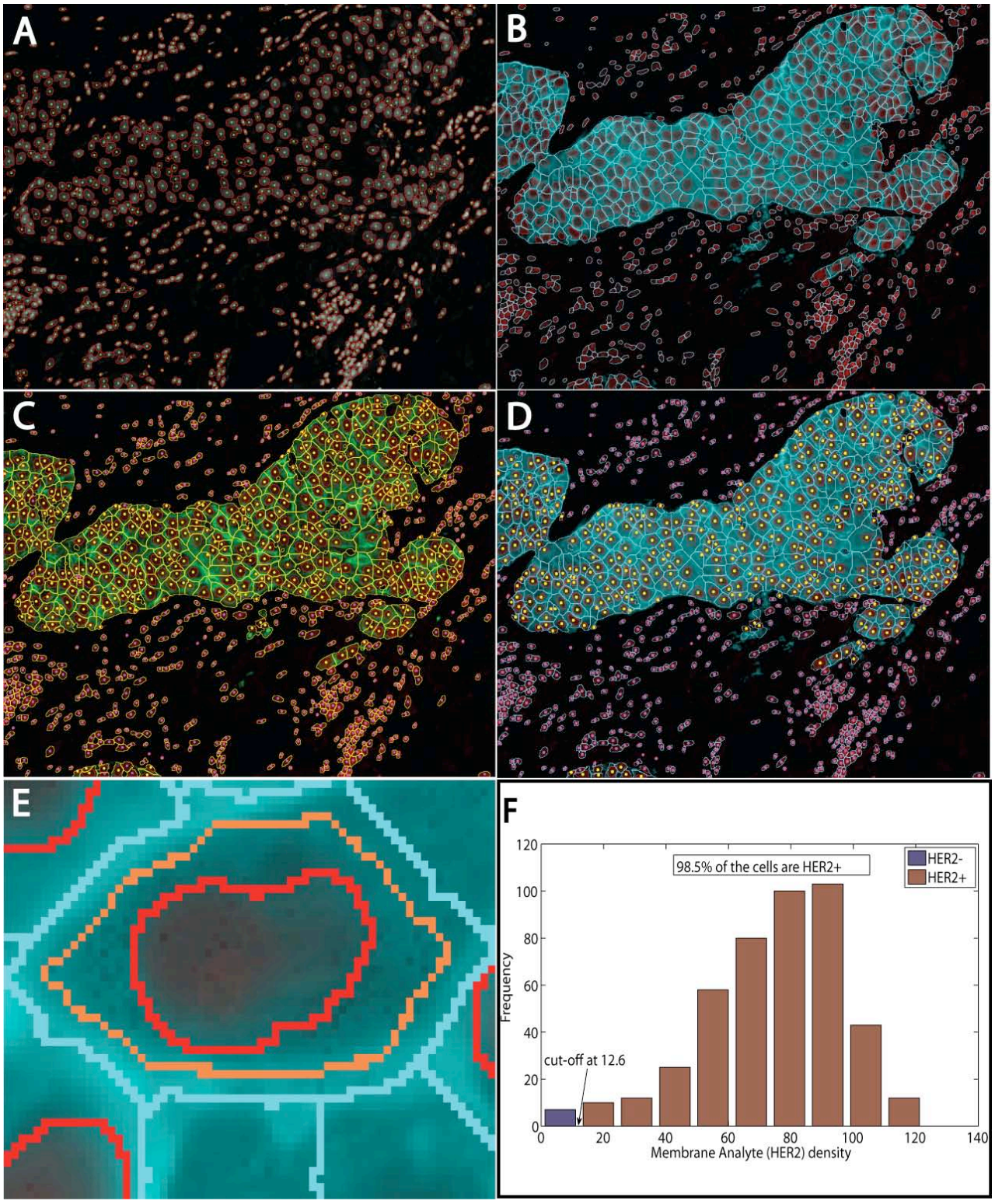

Figure 3: Illustrating the analysis steps for the specimen in Figure 2. (A) Automatic nuclear segmentation results (red outlines) overlaid on the nuclear channel, and green dots indicating seed points. (B) Estimated cytoplasmic domains are shown as blue outlines (Case I). (C) Automated segmentation of the cytokeratin channel (Case II). (D) Composite cell classification results, with yellow dots indicating cells that are cytokeratin and HER2 positive, and purple dots indicating other cells. (E) Close-up view illustrating the regions of interest used to quantify HER2. (F) Histogram summary of the cells in this field that were cytokeratin positive: 98.5% of these cells were also HER2 positive (manual reading 99%).

Task 4. Validate immunohistological analysis system and data obtained
Now that FARSIGHT has been developed for segmenting nuclei and whole breast cancer cells in appropriately stained histological images, we will compare its quantification of ER, PR and HER2 expression in breast cancer specimens that have been previously assessed for these biomarkers by pathologists using IHC-stained slides (for ER, PR, HER2) and by FISH (for HER2). This will be done in the no-cost extension year that we have been granted to complete this project.

Key Research Accomplishments
We have developed robust tissue immunostaining protocols that allow breast cancer specimens to be stained for ER, PR, Ki67 and p-ERK analytes using DAB chromogen along with histochemical staining for nuclei (with hematoxylin) and immunostaining for cytosol (CK) and plasma membrance (E-cad, HER2) by fluorescence.

We have optimized multispectral image capture of specimens immunostained for multiple antigens and structures such that staining for each is separated into distinct channels.

We have incorporated the new Vectra Multispectral Imaging System (CRi) with its Inform software system into our analytical platform.

We have optimized FARSIGHT software for nuclear segmentation based on hematoxylin staining.

We have developed FARSIGHT software for whole cell segmentation; the program segments cells in images stained for cytosolic and/or plasma membrane antigens and compartmentalizes each into nuclear and extranuclear compartments for the purpose of analyte quantification and distribution.

Reportable Outcomes (Publication bibliography)
No publications yet. Two manuscripts have been submitted for peer-review publication' one is being revised following initial review.

Al-Kofahi Y, Lassoued W, Lee WMF, and Roysam B: Improved automatic detection and segmentation of cell nuclei in histopathology images. Trans. Biomed. Eng (under revision).

Al-Kofahi Y, Lassoued W, Grama K, Nath SK, Feldman M, Lee WMF, and Roysam B: Method for cell-based quantification of molecular biomarkers in histopathology specimens from multi-channel images. (submitted).

Personnel receiving pay: William Lee (PI, University of Pennsylvania, Philadelphia, PA)
Badrinath Roysam (co-PI, Rensselaer Polytechnic Institute, Troy, NY)
Wiem Lassoued (grad student/research specialist, University of Pennsylvania, Philadelphia, PA)
Youssef Al-Kofahi (grad student/research specialist, Rensselaer Polytechnic Institute, Troy, NY)
Kedar Grama (grad student/research specialist, Rensselaer Polytechnic Institute, Troy, NY)

Conclusions

Multiple antigens reporting on prognostic/predictive biomarkers, cell signaling and cell fate decisions can be stained for on the same slide in human breast cancer specimens, along with markers of different subcellular compartments in tumor cells.

Following separation of staining for these analytes and subcellular compartment markers into separate channels by multispectral microscopy and spectral unmixing, image analysis can accurately and rapidly segment nuclei and cells in images and associate analytes with the segmented cells and subcellular compartments computationally.